AD-A260 442

Transformations of Gaussian Random Fields and a Test for Independence of a Survival Time from a Covariate

Ian W. McKeague[1]
Department of Statistics
Florida State University
Tallahassee, FL 32306-3033
U.S.A.

A. M. Nikabadze
A. Razmadze Mathematical Institute
Rukhadze St. 1
380093 Tbilisi
Republic of Georgia

Yanqing Sun
Department of Statistics
Florida State University

FSU Technical Report No. M-879
USARO Technical Report No. D-130
AFOSR Technical Report No. 91-275

November, 1992

[1]Research partially supported by Army Research Office Grant DAA03-90-G0103 and by the Air Force Office of Scientific Research under Grant AFOSR91-0048.

DTIC QUALITY INSPECTED 3

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

UNCLASSIFIED MASTER COPY - FOR REPRODUCTION PURPOSES (2)

SECURITY CLASSIFICATION OF THIS PAGE

DTIC ELECTE FEB 1 8 1993 S C D

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY NA	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) FSU Technical Report No. M-879	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARO 27868.24-MA
6a. NAME OF PERFORMING ORGANIZATION Florida State University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U. S. Army Research Office
6c. ADDRESS (City, State, and ZIP Code) Department of Statistics Tallahassee, FL 32306-3033	7b. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U. S. Army Research Office 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAL03-90-G-0103
8c. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
TRANSFORMATIONS OF GAUSSIAN RANDOM FIELDS AND A TEST FOR INDEPENDENCE OF A SURVIVAL TIME FROM A COVARIATE

12. PERSONAL AUTHOR(S)
Ian W. McKeague, A. M. Nikabadze, and Yanqing Sun

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	November, 1992	27

16. SUPPLEMENTARY NOTATION
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Competing risks, ordered alternatives, cumulative incidence function, distribution-free tests, right-censored data, counting processes.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

It has been almost sixty years since Kolmogorov introduced a distribution-free omnibus test for the simple null hypothesis that a distribution function coincides with a given distribution function. Doob subsequently observed that Kolmogorov's approach could be simplified by transforming the empirical process to an empirical process based on uniform random variables. Recent use of more sophisticated transformations has led to the construction of asymptotically distribution-free omnibus tests when unknown parameters are present. The purpose of the present paper is to use the transformation approach to construct an asymptotically distribution-free omnibus test for independence of a survival time from a covariate. The test statistic is obtained from a certain test statistic *process* (indexed by time and covariate), which is shown to converge in distribution to a Brownian sheet. A simulation study is carried out to investigate the finite sample properties of the proposed test and an application to data from the British Medical Research Council's 4th myelomatosis trial is given.

93-03111

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Iam W. McKeague	22b. TELEPHONE (Include Area Code) (904) 644-3218 / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Transformations of Gaussian Random Fields and a Test for Independence of a Survival Time from a Covariate

Ian W. McKeague A. M. Nikabadze Yanqing Sun

Abstract

It has been almost sixty years since Kolmogorov introduced a distribution-free omnibus test for the simple null hypothesis that a distribution function coincides with a given distribution function. Doob subsequently observed that Kolmogorov's approach could be simplified by transforming the empirical process to an empirical process based on uniform random variables. Recent use of more sophisticated transformations has led to the construction of asymptotically distribution-free omnibus tests when unknown parameters are present. The purpose of the present paper is to use the transformation approach to construct an asymptotically distribution-free omnibus test for independence of a survival time from a covariate. The test statistic is obtained from a certain test statistic *process* (indexed by time and covariate), which is shown to converge in distribution to a Brownian sheet. A simulation study is carried out to investigate the finite sample properties of the proposed test and an application to data from the British Medical Research Council's 4th myelomatosis trial is given.

1 Introduction

A standard way of testing for independence of a survival time from a covariate z is to fit Cox's (1972) model for the conditional hazard function, $\lambda(t|z) = \lambda_0(t)e^{\beta z}$, and test whether the regression parameter β is zero. However, this test has limited power because of the restrictive (viz parametric and multiplicative) modeling of the covariate effect.

In this paper we develop an omnibus test that can detect *arbitrary* forms of dependence of a (possibly censored) survival time on a one-dimensional covariate, and which is asymptotically distribution-free. The latter property will be achieved via the transformation method of Doob (1949) and Khmaladze (1981).

We begin by giving some background to the general problem of constructing omnibus tests (i.e. tests consistent against all alternatives) which have the distribution-free property. First consider the simple hypothesis $F = F_0$, where F_0 is specified and the life times $T_1, \cdots, T_n$ are completely observed iid random variables having distribution function F. Let $\hat{F}(t) =$

[1] *AMS 1991 subject classifications.* Primary: 62G10; secondary: 62G20

[2] *Key words and Phrases.* Distribution-free omnibus test, innovation Brownian sheet, counting processes, martingale methods.

$\frac{1}{n}\sum_{i=1}^{n} I(T_i \le t)$ be the empirical distribution function of the T_i's and $\nu_n(t) = \sqrt{n}(\hat{F}(t) - F_0(t))$ the empirical process. Assume that F_0 is continuous. Doob (1949) transformed $\nu_n(t)$ to the uniform empirical process $u_n(r) = \nu_n(F_0^{-1}(r))$, which is an empirical process based on the iid uniform random variables $F_0(T_i)$, $i = 1, \cdots, n$. The distribution of u_n does not depend on F_0 (and it converges weakly to a Brownian bridge), so the distribution of any test statistic that is a functional of u_n is free from F_0. In particular, the Kolmogorov–Smirnov statistic $\sup_x |u_n(x)|$ and the Cramér–von Mises statistic $\int u_n^2(x)\, dx$ are distribution-free.

Next consider the composite null hypothesis $F = F_0(\cdot, \theta)$, where θ is an unknown parameter. The natural extension of the above transformation, $\hat{u}_n(x) = \hat{\nu}_n(F_0^{-1}(x, \hat{\theta}))$, where $\hat{\nu}_n(t) = \sqrt{n}(\hat{F}(t) - F_0(t, \hat{\theta}))$ is the parametric empirical process and $\hat{\theta}$ is an estimator of θ, is unfortunately no longer distribution-free or even asymptotically distribution-free (Durbin, 1973). As a consequence, classical statistics such as $\sup_x |\hat{u}_n(x)|$ or $\int \hat{u}_n^2(x)\, dx$ have limit distributions which depend on F_0. Thus it is necessary to construct a more sophisticated transformation of $\hat{\nu}_n$ that can provide the basis for goodness-of-fit tests, generalizing what the uniform empirical process does in the case of simple hypotheses. Khmaladze (1981) introduced martingale methods to address this problem; see also Nikabadze (1987). The parametric empirical process $\hat{\nu}_n$ converges weakly to some zero-mean Gaussian process ν (Durbin, 1973), so Khmaladze first transformed the process ν to an innovation martingale, which is a Gaussian process with independent increments and covariance function $F_0(s \wedge t, \theta)$ and which preserves the information in ν. Then he transformed the innovation martingale to a standard Brownian motion w. Applying the transformation $\nu \mapsto w$ to $\hat{\nu}_n$, results in a test process that converges weakly to Brownian motion. This leads to an asymptotically distribution-free omnibus test.

Chi-squared tests are widely used for goodness-of-fit testing and for testing independence of two variables in a contingency table analysis. They were first introduced by Pearson (1900) for simple hypotheses $F = F_0$. The chi-squared statistic is formed by dividing part of the real line into cells and comparing the observed and expected frequency in each cell. Fisher (1922, 1924) extended this statistic to handle the presence of an unknown parameter θ in F_0. Chi-squared tests depend on an arbitrary choice of intervals and they only use grouped data. Although chi-squared tests are easy to perform, they are not omnibus (unless the variables are discrete) and are typically less powerful than tests of Kolmogorov–Smirnov or Cramér–von Mises type, which use all the information in the data.

In survival analysis, one is rarely able to observe complete life histories. Important examples occur with right censoring and left truncation (Keiding and Gill, 1990). These examples fit into the general setting of Aalen's (1978) multiplicative intensity model for counting processes. In that setting it is natural to formulate hypotheses in terms of the hazard function $\lambda(t)$ or the cumulative hazard function $\Lambda(t) = \int_0^t \lambda(s)\, ds$, rather than the distribution function F. Andersen et al. (1982) studied tests of the simple hypothesis $\lambda = \lambda_0$ in terms of functionals of $\sqrt{n}(\hat{\Lambda} - \Lambda_0)$, where $\hat{\Lambda}$ is the Nelson–Aalen estimator. Hjort (1990) considered the composite hypothesis $\lambda = \lambda_0(\cdot, \theta)$, with statistics based on functionals of the process $\sqrt{n}(\hat{\Lambda}(t) - \Lambda_0(t, \hat{\theta}))$, where $\hat{\theta}$ is the maximum likelihood estimator of θ. This process converges weakly to a zero-mean Gaussian process under the null hypothesis, and can be used to construct chi-squared tests. Alternatively, an innovation martingale can be found for the limit process and used construct an asymptotically distribution-free omnibus test.

In many applications of survival analysis it is important to consider whether a covariate has some effect upon survival, say through the conditional hazard function $\lambda(t|z) = \lambda(t, z)$. That is, one would like to test the null hypothesis

$$H_0 : \lambda(t, z) \text{ does not depend on the covariate } z,$$

against the general alternative that $\lambda(t, z)$ depends on z. For simplicity, we shall restrict the domain of (t, z) to be the unit square. An omnibus test of H_0 is feasible when the covariate is one-dimensional, such as age at diagnosis, disease duration, etc. Indeed, McKeague and Utikal (1990, subsequently MU) proposed such a test based on the process $X(t, z) = \sqrt{n}(\hat{\mathcal{A}} - \tilde{\mathcal{A}})$, where $\hat{\mathcal{A}}$ is an estimate of the doubly cumulative hazard function $\mathcal{A}(t, z) = \int_0^t\int_0^z \lambda(s, x)\,dx\,ds$, and $\tilde{\mathcal{A}}(t, z) = z\hat{\Lambda}(t)$ is the natural estimate of $\mathcal{A}$ under H_0. They showed that X converges weakly under H_0 to a Gaussian random field of the form

$$m(t, z) = \int_0^t\int_0^z \sqrt{h}\,dW - b(z)\int_0^t\int_0^1 g\,dW, \tag{1.1}$$

where W is a Brownian sheet and h, b, g are certain nonrandom functions. The above stochastic integrals are defined in the L^2-sense, see Wong and Zakai (1974). MU's test was based on the Kolmogorov–Smirnov statistic computed directly from X. However, while asymptotically omnibus, such a test is not asymptotically distribution-free and would require simulation of the process m to find critical values.

We shall construct a transformation J that maps m to its innovation Brownian sheet. An estimated version $\hat{J}$ of J will be obtained by plugging an estimate of h into J (it turns out that J does not involve g and b is known). We then show that $\hat{J}(X)$ converges weakly to Brownian sheet. In this way we obtain an asymptotically distribution-free omnibus test for H_0, with the Kolmogorov–Smirnov statistic computed from $\hat{J}(X)$. No simulation technique is needed to find critical values. The test statistic converges weakly to $\sup|W(t, z)|$. Although an exact formula for the distribution function of $\sup|W(t, z)|$ is not known (only approximations are available, see Adler (1991)), it is straightforward to carry out a single Monte Carlo experiment to evaluate it quite accurately. Thus, our test avoids difficulties arising from simulating the null distribution for each particular problem.

The paper is organized as follows. In Section 2, we construct the transformation J. In Section 3, we introduce the estimate $\hat{J}$ and define the test statistic. Results of a simulation study are reported in Section 4. In Section 5, the test is applied to a set of data from the British Medical Research Council's (1984) 4th myelomatosis trial. Properties of the test are proved in Section 6. Various lemmas needed through the paper are collected in an appendix.

2 Transformation of m to Brownian Sheet

In this section we construct our transformation J of the Gaussian random field m in (1.1) to Brownian sheet. Such a transformation is likely to have further applications in nonparametric statistics beyond our test for independence—in any setting where a test process converges weakly to a process of the form (1.1): e.g. in testing whether $\lambda(t, z)$ is independent of t (i.e. the roles of t and z are reversed), or testing whether a pure jump process on a finite state space is a semi-Markov process, see MU (Section 1.2). Of course, it is usually necessary to estimate J and how that is done will depend on the particular application.

We begin with a key proposition showing that the law of a Brownian sheet W is preserved under a shift of W by a certain functional of W.

Proposition 2.1 *Let $k \in L^2([0,1]^2)$ satisfy $\int_u^1 k^2(s,v)\,dv > 0$ a.e. $[ds]$ for $u < 1$, and let W be a Brownian sheet. Then*

$$B(t,z) = W(t,z) - \int_0^z \left[\int_0^t \int_u^1 \frac{k(s,x)k(s,u)}{\int_u^1 k^2(s,v)\,dv}\, dW(s,x) \right] du \tag{2.2}$$

is a Brownian sheet on $[0,1]^2$. The relation (2.2) is invertible:

$$W(t,z) = B(t,z) + \int_0^z \left[\int_0^t \int_u^1 \frac{k(s,x)k(s,u)}{\int_x^1 k^2(s,v)\,dv}\, dB(s,x) \right] du. \tag{2.3}$$

Proof Let

$$a(t,u;s,x) = \frac{k(s,x)k(s,u)I(x \geq u)I(s \leq t)}{\int_u^1 k^2(s,v)\,dv}.$$

Then

$$B(t,z) = W(t,z) - \int_0^z \left[\int_0^1 \int_0^1 a(t,u;s,x)\, dW(s,x) \right] du.$$

Notice that B is a Gaussian random field, so we only need to inspect its covariance function. For $(t',z') \in [0,1]^2$,

$$\begin{aligned}
\operatorname{cov}(B(t,z), B(t',z')) &= (t \wedge t')(z \wedge z') \\
&\quad - \int_0^{z'} \left[\int_0^t \int_0^z a(t',u';s,x)\, ds\, dx \right] du' - \int_0^z \left[\int_0^{t'} \int_0^{z'} a(t,u;s,x)\, ds\, dx \right] du \\
&\quad + \int_0^z \int_0^{z'} \left[\int_0^1 \int_0^1 a(t,u;s,x) a(t',u';s,x)\, ds\, dx \right] du\, du' \\
&= (t \wedge t')(z \wedge z') \\
&\quad + \int_0^z \int_0^{z'} \left[\int_0^1 \int_0^1 a(t,u;s,x) a(t',u',s,x)\, ds\, dx \right. \\
&\quad \left. - \int_0^1 a(t,u;s,u') I(s \leq t')\, ds - \int_0^1 a(t',u';s,u) I(s \leq t)\, ds \right] du'\, du.
\end{aligned}$$

Since

$$\begin{aligned}
&\int_0^1 a(t,u;s,x) a(t',u';s,x)\, dx \\
&\quad = \frac{k(s,u)I(s \leq t)k(s,u')I(s \leq t') \int_0^1 k^2(s,x) I(x \geq u \vee u')\, dx}{\int_u^1 k^2(s,v)\, dv \int_{u'}^1 k^2(s,v)\, dv} \\
&\quad = \frac{k(s,u')k(s,u)I(u' \geq u)I(s \leq t)I(s \leq t')}{\int_u^1 k^2(s,v)\, dv} \\
&\qquad + \frac{k(s,u')k(s,u)I(u \geq u')I(s \leq t)I(s \leq t')}{\int_{u'}^1 k^2(s,v)\, dv} \\
&\quad = a(t,u;s,u')I(s \leq t') + a(t',u',s,u)I(s \leq t),
\end{aligned}$$

for almost all $(u,u',s) \in [0,1]^3$, we have that B is a Brownian sheet. It can be verified immediately that (2.3) is the inverse of (2.2). $\square$

A Brownian motion $w(t)$ is called an *innovation process* of a process $\xi(t)$ if w carries the same "information" as the process ξ, i.e. the σ-fields $\mathcal{F}_t^w$ and $\mathcal{F}_t^\xi$ generated by w and ξ up to each time t coincide, see Liptser and Shiryayev (1977, p. 260). For our purposes, the appropriate extension of this definition to a two-parameter process $\xi(t,z)$ and a Brownian sheet $B(t,z)$ is made by requiring $\mathcal{F}_t^B = \mathcal{F}_t^\xi$, where $\mathcal{F}_t^\xi = \sigma\{\xi(s,z); z \in [0,1], s \in [0,t]\}$ and $\mathcal{F}_t^B$ is similarly defined. Note that $\mathcal{F}_t^\xi$ represents the information about $\xi(s,z)$ at all values of z and all $s \le t$.

We now give the main result of this section, providing an innovation Brownian sheet for the process m in (1.1).

Theorem 2.1 *Suppose that $h : [0,1]^2 \to I\!R$ is a bounded positive measurable function which is bounded away from zero, $b : [0,1] \to I\!R$ is differentiable with square integrable derivative, $\int_z^1 (b'(x))^2\,dx > 0$, $z \in [0,1)$ and $g \in L^2([0,1]^2)$. Then*

$$B(t,z) = \int_0^t\!\!\int_0^z h^{-\frac12}\,dm - \int_0^t\!\!\int_0^1 \left[\int_0^{u\wedge z} h^{-\frac12}(s,u)Q(s,u,x)\,dx\right] dm(s,u) \qquad (2.4)$$

is an innovation Brownian sheet of the process m, where

$$Q(s,u,x) = \frac{h^{-\frac12}(s,u)b'(u)h^{-\frac12}(s,x)b'(x)}{\int_x^1 h^{-1}(s,v)(b'(v))^2\,dv}.$$

Proof Notice that

$$\int_0^1\!\!\int_0^{u\wedge z} Q(s,u,x)h^{-\frac12}(s,u)b'(u)\,dx\,du = \int_0^z h^{-\frac12}(s,x)b'(x)\,dx.$$

Let

$$U(t) = \int_0^t\!\!\int_0^1 g\,dW.$$

Substituting m into (2.4) we get

$$\begin{aligned}
B(t,z) &= W(t,z) - \int_0^t \left[\int_0^z h^{-\frac12}(s,x)b'(x)\,dx\right] U(ds) \\
&\quad - \int_0^t\!\!\int_0^1 \left[\int_0^{u\wedge z} Q(s,u,x)\,dx\right] dW(s,u) \\
&\qquad + \int_0^t\!\!\int_0^1 \left[\int_0^{u\wedge z} Q(s,u,x)\,dx\right] h^{-\frac12}(s,u)b'(u)\,du\,U(ds) \\
&= W(t,z) - \int_0^t\!\!\int_0^1 g(s,y)\left[\int_0^z h^{-\frac12}(s,x)b'(x)\,dx\right] dW(s,y) \\
&\quad - \int_0^z \left[\int_0^t\!\!\int_x^1 Q(s,y,x)\,dW(s,y)\right] dx \\
&\quad + \int_0^t\!\!\int_0^1 \left[g(s,y)\int_0^1\!\!\int_0^{u\wedge z} Q(s,u,x)h^{-\frac12}(s,u)b'(u)\,dx\,du\right] dW(s,y) \\
&= W(t,z) - \int_0^z \left[\int_0^t\!\!\int_x^1 Q(s,y,x)\,dW(s,y)\right] dx. \qquad (2.5)
\end{aligned}$$

This is a Brownian sheet by Lemma 2.1 with $k(s,x) = h^{-\frac12}(s,x)b'(x)$. The innovation assertion can be easily obtained by (2.4), (2.5) and application of Proposition 2.1. □

We shall use the notation J for the transformation $\xi \longmapsto J(\xi)$, where ξ is a random field and $J(\xi)$ is defined by the right side of (2.4) with m replaced by ξ. The domain of J is composed of random fields ξ for which the stochastic integrals in $J(\xi)$ exist in the L^2-sense. Theorem 2.1 shows that $J(m)$ is a Brownian sheet.

3 The Test Procedure

In this section we first describe the counting process framework for our problem and formally define $\hat{A}$ and $\bar{A}$. Then we show that the transformation J given above asymptotically transforms $X = \sqrt{n}(\hat{A} - A)$ to a Brownian sheet. This is done via the continuous mapping theorem. Finally, we construct an estimate $\hat{J}$ of J and show that $\hat{J}(X)$ converges weakly to a Brownian sheet. This will complete the construction of our test.

3.1 The Estimators $\hat{A}$ and $\bar{A}$

Let $\mathbf{N}(t) = (N_1(t), \cdots, N_n(t))$, $t \in [0,1]$, be a multivariate counting process with respect to a right-continuous filtration $(\mathcal{F}_t)$, i.e., $\mathbf{N}$ is adapted to the filtration and has components N_i which are right-continuous step functions, zero at time zero, with jumps of size $+1$ such that no two components jump simultaneously. Assume that N_i has intensity

$$\lambda_i(t) = Y_i(t)\lambda(t, Z_i(t)),$$

where Y_i is a predictable $\{0, 1\}$-valued process, indicating that the ith individual is at risk when $Y_i(t) = 1$, and Z_i is a predictable $[0,1]$-valued covariate process. The function $\lambda(t,z)$ represents the failure rate for an individual at time t with covariate $Z_i(t) = z$. We assume throughout that (N_i, Y_i, Z_i), $i = 1, \cdots, n$ are iid replicates of an underlying triple (N, Y, Z). Let $F(s,x) = P(Z_s \le x, Y_s = 1)$, and assume that for each $s \in [0,1]$, $F(s,\cdot)$ is absolutely continuous on $[0,1]$ with subdensity $f(s,\cdot)$. The functions b, h, g in (1.1) are given by $b(z) = z$, $h = \lambda/f$ and $g = \sqrt{\lambda \cdot f}$. The transformation J will only be used with these b and h from now on. We assume that f and λ are Lipschitz, of bounded variation, and bounded away from zero.

Consider d_n equal width covariate strata $\mathcal{I}_r = [x_{r-1}, x_r)$, $r = 1, \ldots, d_n$, where $x_r = rw_n$, and $w_n = 1/d_n$ is the stratum width, and let $\mathcal{I}_z = \mathcal{I}_r$ for $z \in \mathcal{I}_r$. Then, as in MU, define

$$\hat{A}(t, z) = \int_0^z \int_0^t \frac{N^{(n)}(ds, x)}{Y^{(n)}(s, x)}\, dx, \tag{3.6}$$

where $N^{(n)}(t, z) = \sum_{i=1}^n \int_0^t I(Z_i(s) \in \mathcal{I}_z)\, dN_i(s)$ is the number of z-specific failures observed up to time t, and $Y^{(n)}(t, z) = \sum_{i=1}^n I(Z_i(t) \in \mathcal{I}_z) Y_i(t)$ is the size of the z-specific risk set at time t. The estimator $\bar{A}$ does not involve stratification of the covariate and can be obtained by setting $w_n = 1$ in $\hat{A}$. In (3.6) and throughout the paper, we use the convention $1/0 \equiv 0$.

3.2 A Continuous Version of J

We now introduce a version $\bar{J}$ of J that is defined on a suitably large function space and is continuous on a subspace supporting m, so the continuous mapping theorem is applicable.

Let $D_2 = D_2([0,1]^2)$ be the extension of the usual Skorohod space to functions on $[0,1]^2$, see Neuhaus (1971). Let BV_2 denote the subspace of functions $\xi \in D_2$ for which $\xi, \xi(0,\cdot), \xi(\cdot,0)$ have bounded variation, and let C_2 denote the space of continuous functions on $[0,1]^2$. Equip D_2 and C_2 with the uniform norm.

For $\xi \in C_2 \cup BV_2$ and $(t,z) \in [0,1] \times [0,\rho]$, with $0 < \rho < 1$, define

$$\bar{J}(\xi)(t,z) = \int_0^t\int_0^z f_1(s,x)\, d\xi(s,x) - \int_0^t\int_0^1 f_2(s,u,z)\, d\xi(s,u), \tag{3.7}$$

where the integrals are considered to be weak integrals (Hildebrandt, 1963), and

$$\begin{aligned} f_1(s,x) &= h^{-\frac{1}{2}}(s,x), \\ f_2(s,u,z) &= h^{-1}(s,u)\int_0^{z\wedge u} \frac{h^{-\frac{1}{2}}(s,x)}{\int_x^1 h^{-1}(s,v)\,dv}\,dx. \end{aligned}$$

The upper bound ρ on the domain of z is used to keep the the denominator in f_2 bounded away from zero. In practice ρ would be taken close to 1 (the end of the range of covariate values). Note that $\bar{J}$ is a well-defined map from $C_2 \cup BV_2$ into $D_2([0,1]\times[0,\rho])$ since Lemma 1 shows that h inherits the properties of f, λ; and Lemma 2 ensures the existence of the weak integrals when $\xi \in C_2$. We have included BV_2 in the domain of J because the paths of X belong to BV_2, but not to C_2.

Theorem 3.1 *Suppose that $w_n \asymp n^{-\alpha}$, where $\frac{1}{2} < \alpha < 1$. Then, under H_0, $\bar{J}(X)$ converges weakly to a Brownian sheet in $D_2([0,1]\times[0,\rho])$.*

Proof Properties of f_1, f_2 obtained via Lemma 2 can be used to show that J is continuous as a map from C_2 into $D_2([0,1]\times[0,\rho])$. In particular, we use the property that $f_2(\cdot,\cdot,z)$ has bounded variation uniformly in z, $0 < z < \rho$. MU (Theorem 4.1) gives that X converges weakly in D_2 to m, where m is defined by (1.1) with $b(z) = z$. Thus, since the sample paths of m belong to C_2 a.s., the continuous mapping theorem (Billingsley, 1968) gives $\bar{J}(X) \xrightarrow{\mathcal{D}} \bar{J}(m)$ in $D_2([0,1]\times[0,\rho])$. The processes $\bar{J}(m)$ and $J(m)$ have continuous sample paths and, by Lemma 3, they agree a.s. at each fixed (t,z), so they are indistinguishable. Theorem 2.1 (with $b(z) = z$) implies that $J(m)$, and hence $\bar{J}(m)$, is a Brownian sheet. □

3.3 Estimating the Transformation

In order to use the above result to build a test statistic, we need to estimate the unknown function in $\bar{J}$, namely h. First consider the kernel estimator $\hat{h}$ suggested by MU:

$$\hat{h}(t,z) = \frac{1}{b_n^2}\int_0^1\int_0^1 K\left(\frac{t-s}{b_n}\right)K\left(\frac{z-x}{b_n}\right) d\hat{H}(s,x),$$

where b_n is bandwidth parameter, K is a Lipschitz nonnegative kernel function with compact support and integral 1, and

$$\hat{H}(t,z) = nw_n \int_0^z\int_0^t \frac{X^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}\,dx$$

is an estimator of $H(t,z) = \int_0^t\int_0^z h(s,x)\,ds\,dx$.

We will need to apply methods from stochastic calculus to various martingale integrals involving $\hat{h}$, which is possible provided that $\hat{h}(\cdot,z)$ is an $\mathcal{F}_t$-predictable process for each fixed z. Since $\hat{h}(\cdot,z)$ is continuous, it is enough that it be adapted to the filtration $\mathcal{F}_t$. Thus, we shall use a kernel function K having nonnegative (as well as compact) support.

The estimated transformation $\hat{J}$ is defined by inserting a truncated version $\check{h}$ of $\hat{h}$ in place of h in J, where $\check{h}$ is given by

$$\check{h}(t,z) = \hat{h}(t,z)I(c_n^{-1} < \hat{h}(t,z) < c_n),$$

$c_n > 0$. Note that $\hat{J}(X)$ is well-defined since the paths of X belong to BV_2.

3.4 The Test Statistic

If we show that $\hat{J}(X)$ converges weakly to a Brownian sheet, then our test for H_0 can be based on the Kolmogorov–Smirnov statistic

$$S = \sup_{0\le t\le 1, 0\le z\le \rho} |\hat{J}(\sqrt{n}(\hat{A} - A))(t,z)|$$

with P-values calculated from the distribution of $\sup_{0\le t\le 1, 0\le z\le \rho} |W(t,z)|$. For that purpose we restrict the choice of w_n, b_n, c_n as follows:

Condition 3.1

$$w_n \asymp n^{-\alpha} \quad \text{where} \quad \tfrac{9}{17} < \alpha < 1,$$

$$b_n \asymp n^{-\beta} \quad \text{where} \quad 0 < \beta < \tfrac{3}{8}(1-\alpha),$$

$$c_n \asymp (\log n)^{\gamma} \quad \text{where} \quad \gamma > 0.$$

This condition is satisfied, for example, by $w_n \asymp n^{-\frac{5}{9}}$, $b_n \asymp n^{-\frac{1}{6}}$.

Theorem 3.2 *Under H_0, $\hat{J}(X)$ converges weakly to a Brownian sheet in $D_2([0,1] \times [0,\rho])$.*

Our final result shows that the test based on S is omnibus, consistent against any departure from the null hypothesis H_0.

Theorem 3.3 *The test based on S is consistent against the general alternative that $\lambda(t,z)$ depends on z, for (t,z) in the domain $[0,1] \times [0,\rho]$.*

4 A Simulation Study

We have carried out a limited simulation study to assess the performance of the proposed test. We considered the Kolmogorov–Smirnov statistic S with the supremum taken over $[0,1] \times [0,.9]$, i.e. $\rho = .9$. The covariate was taken to be uniformly distributed over $[0,1]$. The censoring was simple right censoring, independent of the failure time, and exponentially distributed with parameter adjusted to give a prescribed percentage 5% (low), 40% (moderate), and 75% (heavy) of censored observations before the end of follow-up. The covariate strata were arranged to contain equal numbers of observations. For sample size $n = 180$,

Table 1: Observed levels and powers of the test for independence of a survival time from a covariate with nominal level 5%.

$\lambda(t,z)$	sample size	censoring 5%	40%	75%
1	180	.038	.021	.000
	300	.042	.039	.003
	500	.073	.058	.001
e^z	180	.068	.025	.007
	300	.184	.116	.014
	500	.525	.38	.069
e^{2z}	180	.128	.053	.008
	300	.448	.279	.032
	500	.872	.708	.182

300 and 500, the number of strata d_n was taken to be 12, 15 and 20, resulting in 15, 20, 25 covariate values in each stratum. The corresponding bandwidths b_n were .29, .26, .22, and the kernel function K was taken to be the indicator of $[0,1]$.

The survival times were generated using the Cox model $\lambda(t,z) = e^{\beta_0 z}$, for $\beta_0 = 0$ (null hypothesis), 1 and 2 (alternative hypotheses). Table 4 gives observed levels and powers of the test at a nominal (asymptotic) level 5%, with each entry based on 1000 samples. In order to obtain the asymptotic 5% critical level for our test (i.e., the 95th percentile of $\sup_{0\le t\le 1, 0\le z\le .9} |W(t,z)|$), we generated 10.000 replicates of the Brownian sheet evaluated on a grid defined by 300 equally spaced points on each axis. Plots of the density and distribution function of the supremum of the absolute value of the Brownian sheet over $[0,1]\times[0,.9]$ are given in Figure 1. The 5% critical level was found to be 2.28.

The simulation results show that the observed levels are close to their nominal 5% values when the sample size is at least 180 and censoring is light or moderate. The power reaches 52% at sample size 500 and low censoring, when the alternative is $\lambda(t,z) = e^z$. It exceeds 70% at sample size 500 when the alternative is $\lambda(t,z) = e^{2z}$ and censoring is moderate.

In Figures 2–4 we give plots of the observed densities (each based on 1000 samples) of the Kolomogorov-Smirnov statistic S under the null and alternative hypotheses. They are compared with the density of $\sup_{0\le t\le 1, 0\le z\le .9} |W(t,z)|$. When the sample size is at least 300 and the censoring is light or moderate, the observed densities agree well with their theoretical limit (see Figure 2). Under the alternatives $\lambda(t,z) = e^z$ and e^{2z}, when the sample size is at least 500 and the censoring is light or moderate, the two curves are quite separate, giving some idea of the power of the test (see Figures 3 and 4). In Figure 5, we give perspective plots of a realization of Brownian sheet and a realization of the test process $\hat{J}(X)$ (with $\lambda(t,z) = 1$, sample size 500 and light censoring). As expected, these plots are qualitatively very similar to one another.

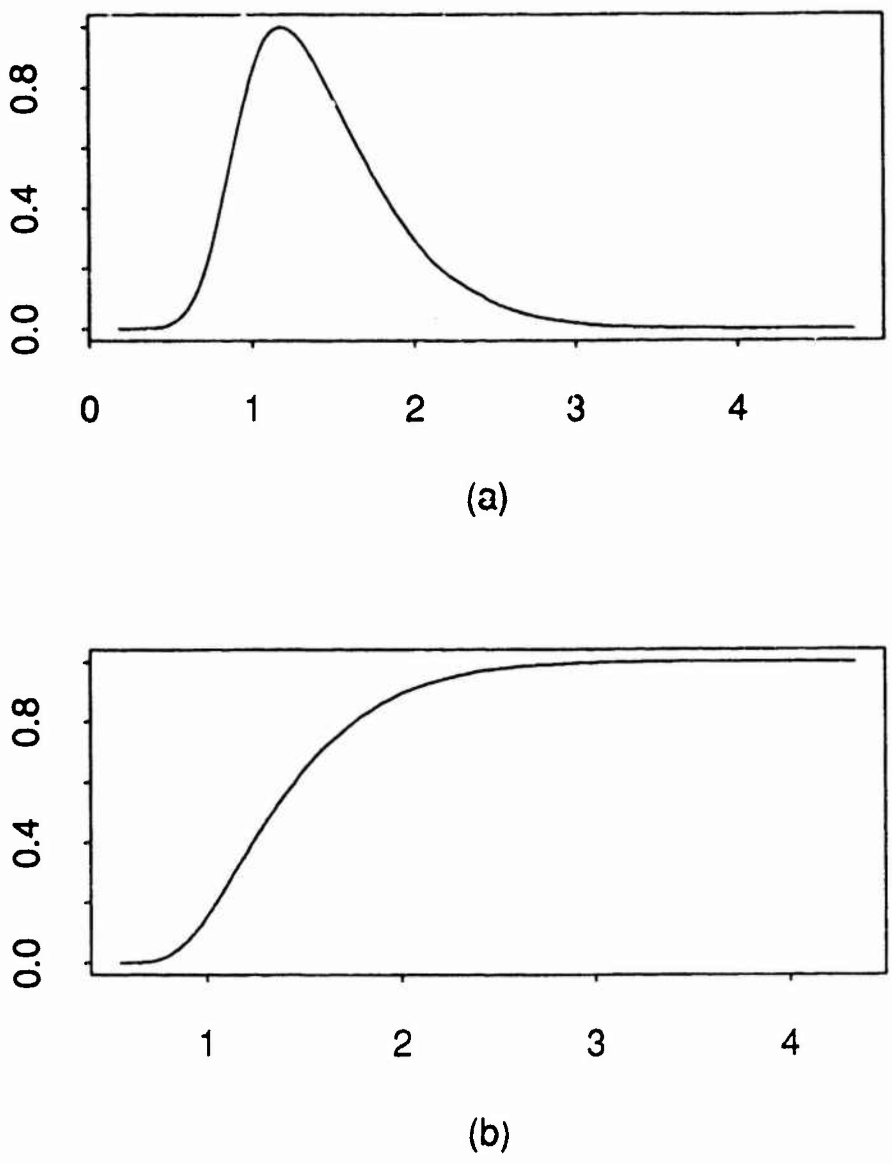

Figure 1: (a) The density and (b) the distribution function of sup $|W(t,z)|$, $(t,z) \in [0,1] \times [0,.9]$.

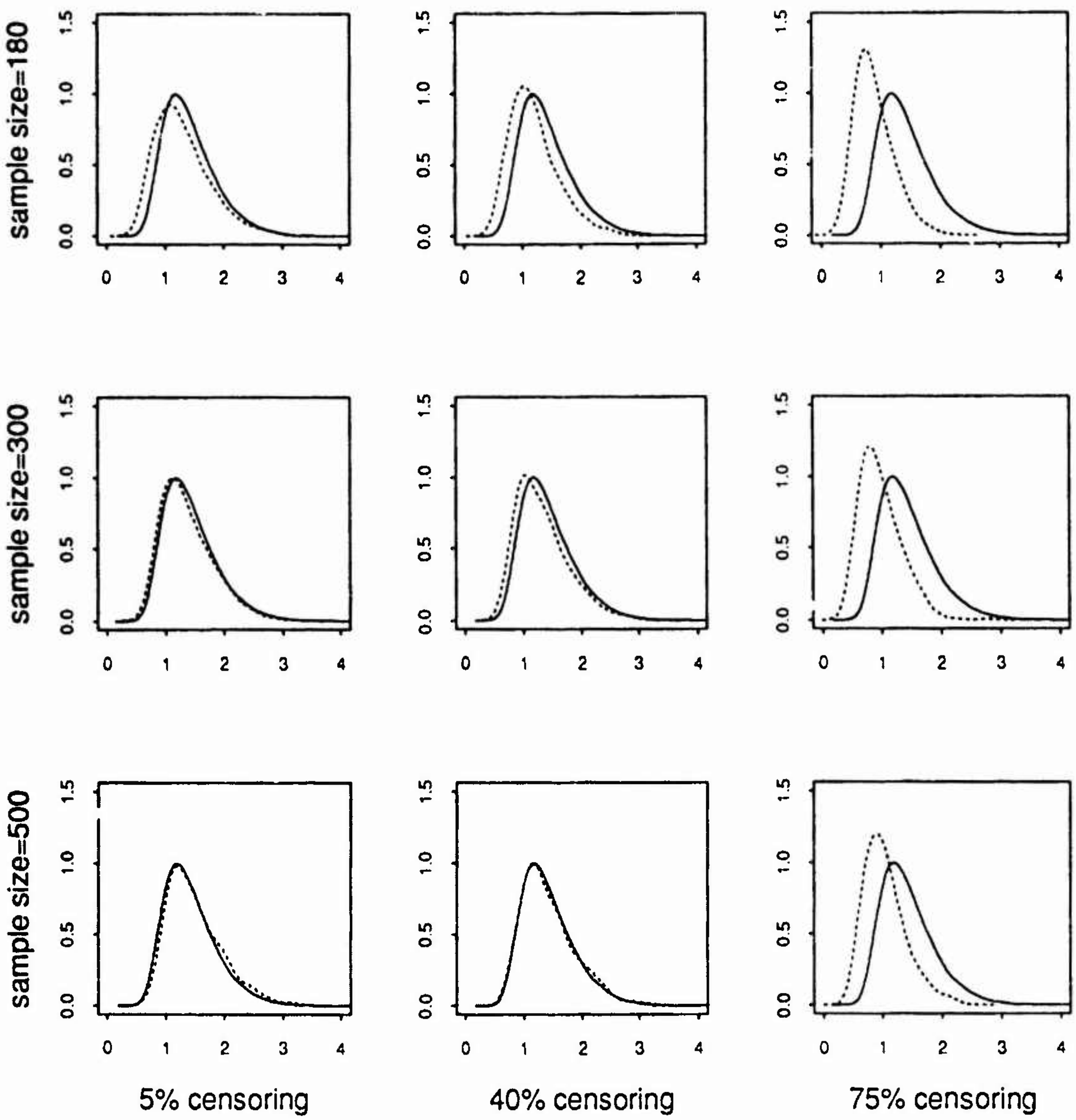

Figure 2: Observed density (dotted line) of the Kolmogorov–Smirnov statistic S under the null hypothesis $\lambda(t,z)=1$ compared with the density (solid line) of $\sup|W(t,z)|$.

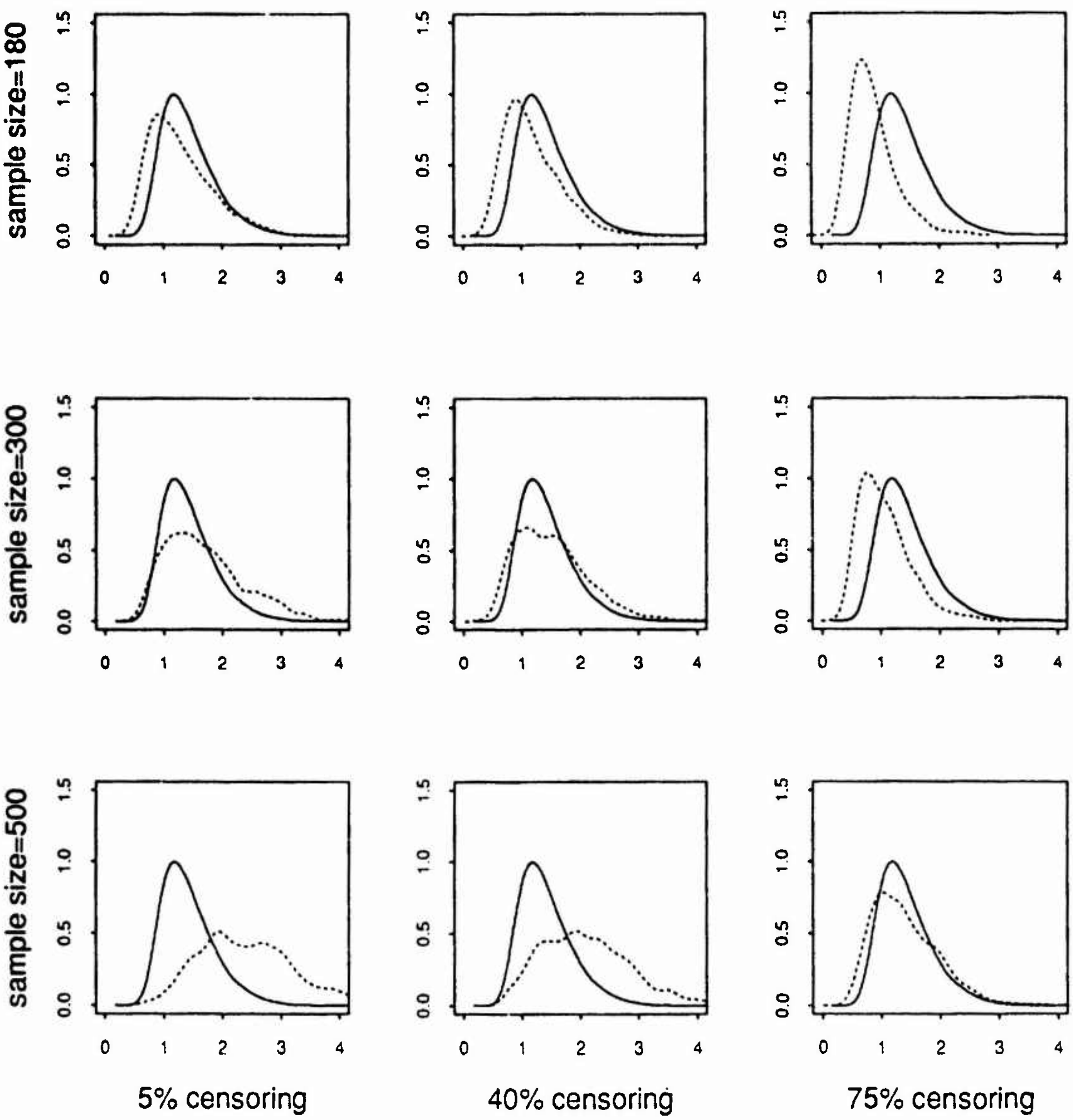

Figure 3: Same as Figure 2, but for the alternative hypothesis $\lambda(t, z) = e^z$.

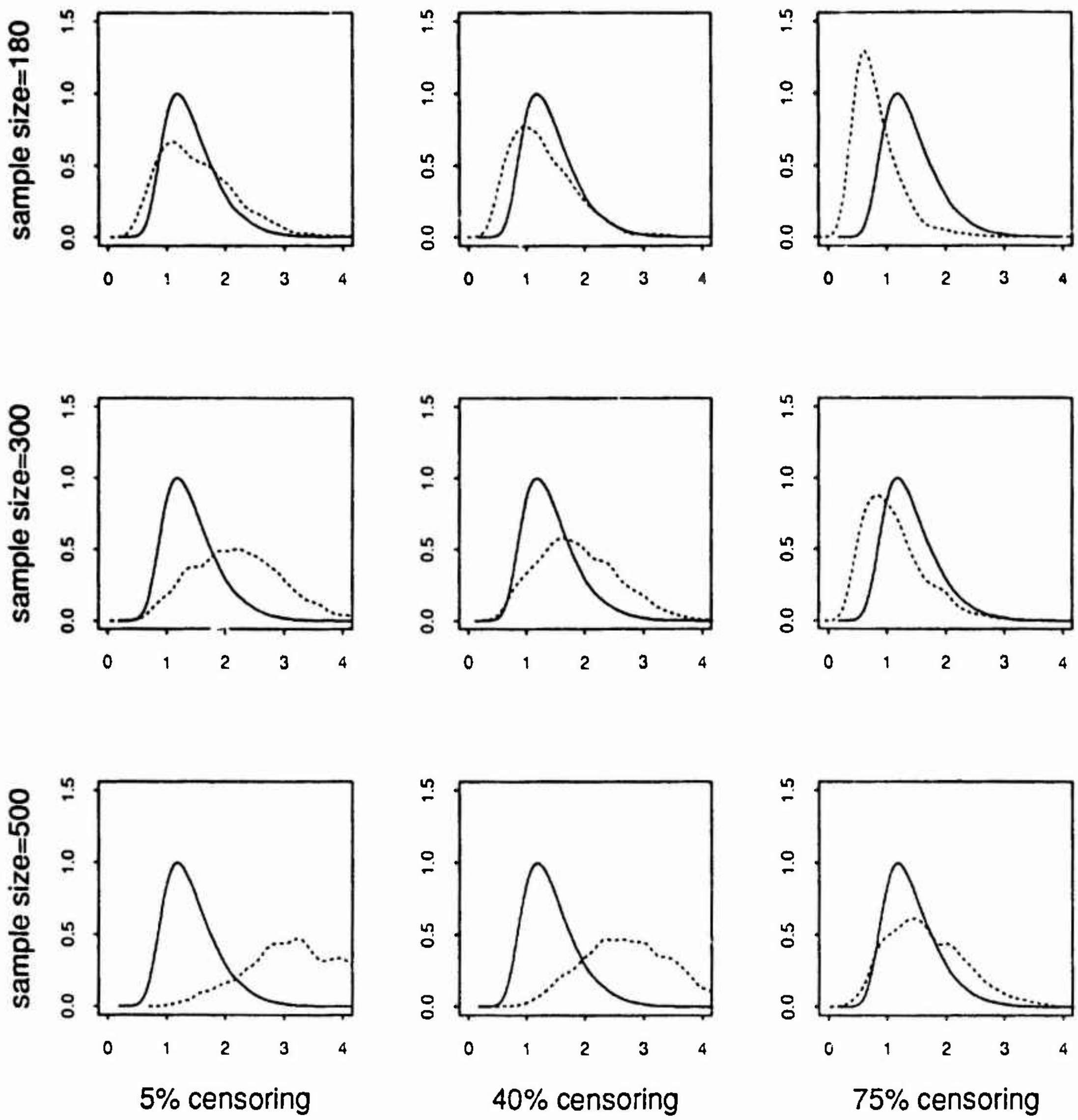

Figure 4: Same as Figure 2, but for the alternative hypothesis $\lambda(t, z) = e^{2z}$.

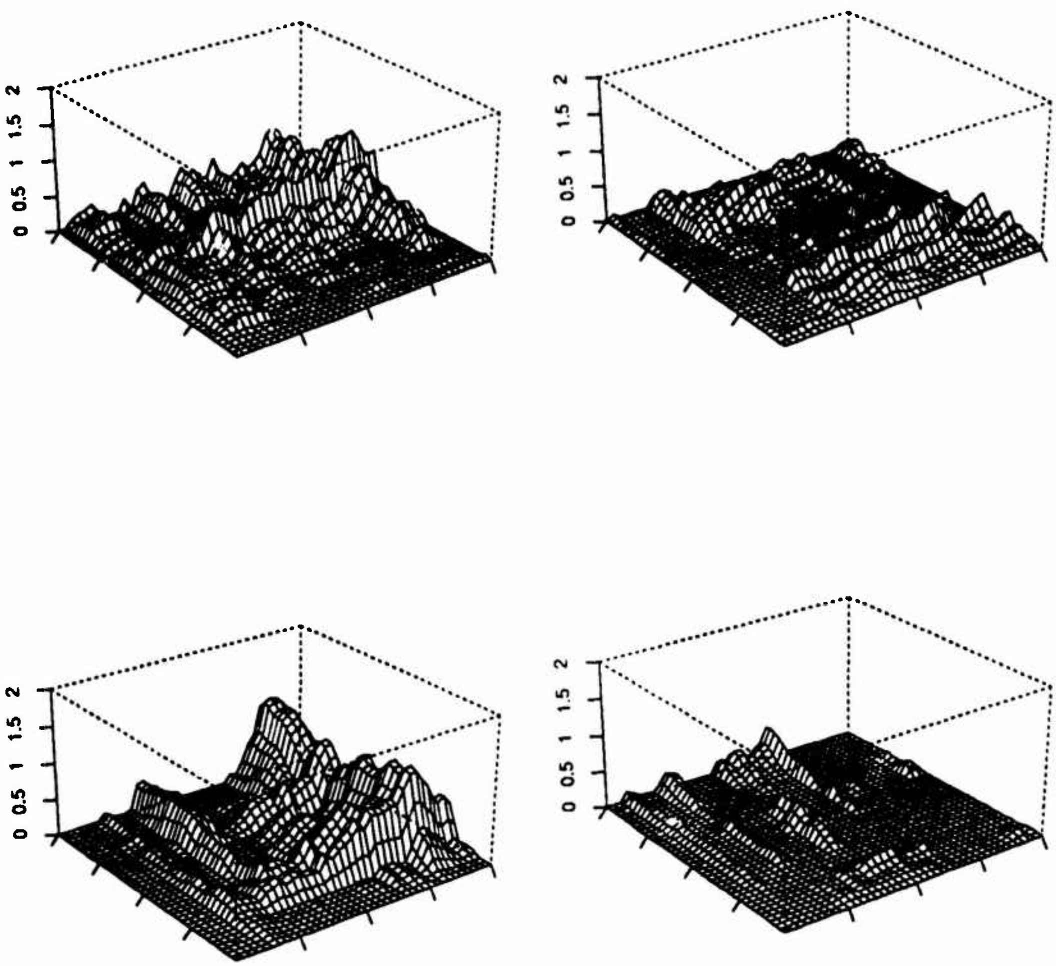

Figure 5: A single realization of Brownian sheet (first row), and a single realization of $\hat{J}(X)$ (second row) from simulated data ($n = 500$, light censoring) under the null hypothesis $\lambda(t, z) = 1$. Positive parts are on the left, negative parts on the right.

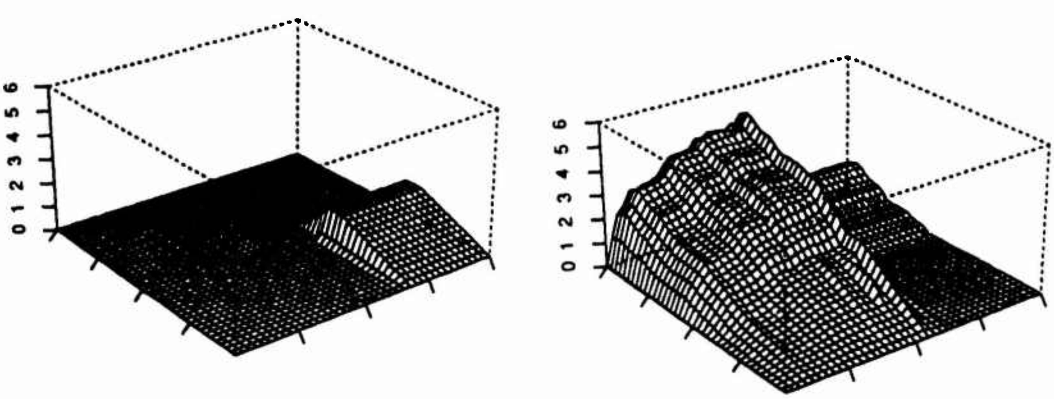

Figure 6: The test process $\hat{J}(X)$ for the BMRC data (positive part on the left, negative part on the right).

5 Application to Myelomatosis Data

We applied our test to a set of data from the British Medical Research Council's (BMRC) (1984) 4th myelomatosis trial. The data set contains records for 495 patients, including censoring indicator, serum β_2 microglobulin (at presentation) and survival time (in days).

Many studies (e.g., Cuzick et al. (1985)) have suggested that serum β_2 microglobulin has a strong effect on survival, at least in the first two years of follow-up. In our analysis of the data, we shall ignore all covariates except for serum β_2 microglobulin, which is taken on a $\log_{10}$ scale normalized to the interval $[0,1]$. The end of follow-up is taken to be 2000 days, before which 3% of the observations are censored. 81 patients were still at risk at the end of follow-up. The survival time is divided by 2000 to normalize it to the interval $[0,1]$. The covariate interval $[0,1]$ is divided so that each stratum contains 20 covariate values except for the last stratum. We used $\rho = .9$, as in the simulation study.

We have plotted the test process $\hat{J}(X)$ in Figure 6. The magnitude of the negative part of $\hat{J}(X)$ shows strong departure from a Brownian sheet, cf. Figure 5. The statistic S was found to be 5.335, which is highly significant. Our analysis confirms the strong influence of a patient's serum β_2 microglobulin on survival.

6 Proofs

In this section we prove Theorems 3.2 and 3.3. We begin by introducing some notation. Let M_i denote the $\mathcal{F}_t$-martingale $M_i(t) = N_i(t) - \int_0^t \lambda_i(s)\,ds$, and set

$$\begin{aligned} M^{(n)}(t,z) &= \sum_{i=1}^{n} \int_0^t I(Z_i(s) \in \mathcal{I}_z)\,dM_i(s) \\ \lambda^{(n)}(t,z) &= \sum_{i=1}^{n} I(Z_i(s) \in \mathcal{I}_z) Y_i(t) \lambda(t, Z_i(t)). \end{aligned}$$

For a process $\xi(t,z)$, set $\xi_r(t) = \xi(t, x_r)$ where $x_r = rw_n$, $r = 1, \ldots, d_n$.

We shall have frequent use for the following bounds from MU (Lemma 1):

$$\sup_{s,x,n} E\left[\frac{nw_n}{Y^{(n)}(s,x)}\right]^k < \infty, \quad \text{for any positive integer } k, \tag{6.8}$$

$$\sup_{s,x} P(Y^{(n)}(s,x) = 0) \leq e^{-Cnw_n}, \quad \text{for some } C > 0. \tag{6.9}$$

Proof of Theorem 3.2 By Theorem 3.1, it is sufficient to show that under H_0,

$$\|(\bar{J} - \hat{J})(X)\| \xrightarrow{P} 0,$$

where $\|\cdot\|$ is the supremum norm on $D_2([0,1] \times [0,\rho])$. This will be done in the following two steps:

$$\left\| \int\!\!\int_0^{\cdot} \bar{f}_1\, dX \right\| \xrightarrow{P} 0, \tag{6.10}$$

$$\left\| \int\!\!\int_0^1 \bar{f}_2(s,u,\cdot)\, dX(s,u) \right\| \xrightarrow{P} 0. \tag{6.11}$$

where $\bar{f}_i = \check{f}_i - f_i$, and $\check{f}_i$ is obtained by inserting $\check{h}$ in place of h in f_i, $i = 1, 2$.

Step 1 By the decomposition of X given in MU (proof of Theorem 4.1),

$$\begin{aligned}\int_0^t\!\!\int_0^z \bar{f}_1 dX &= \sqrt{n}\int_0^z\!\!\int_0^t \bar{f}_1(s,x)\frac{M^{(n)}(ds,x)}{Y^{(n)}(s,x)}\,dx \\ &\quad -\sqrt{n}\int_0^z\int_0^t \bar{f}_1(s,x)\frac{d\bar{M}^{(n)}(s)}{\bar{Y}^{(n)}(s)}\,dx \\ &\quad +\int_0^t\!\!\int_0^z \bar{f}_1(s,x)\sqrt{n}\Big(\frac{\lambda^{(n)}(s,x)}{Y^{(n)}(s,x)} - \lambda(s,x)\Big)\,dx\,ds \\ &\quad +\sqrt{n}\int_0^t\!\!\int_0^z \bar{f}_1(s,x)\lambda(s)I(\bar{Y}^{(n)}(s)=0)\,dx\,ds,\end{aligned} \tag{6.12}$$

where $\bar{M}^{(n)}$, $\bar{Y}^{(n)}$ are defined by setting $\mathcal{I}_z = [0,1]$ in $M^{(n)}$, $Y^{(n)}$, respectively. We denote the four terms in the above decomposition by I_1, I_2, I_3, I_4, respectively. Since K is continuous and has nonnegative support, we have that $\check{h}(\cdot,x)$, and therefore $\bar{f}_1(\cdot,x)$, is $\mathcal{F}_t$-predictable. Thus the stochastic integrals involved in I_1 and I_2 are square integrable martingales. Now $\|I_1\|$ is bounded by

$$\sup_t \eta(t) + \sqrt{n}\sup_{\substack{t,r\\ z\in\mathcal{I}_r}}\left|\int_{\mathcal{I}_{r-}}^z\int_0^t \bar{f}_1\frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}\,dx\right| \tag{6.13}$$

where

$$\eta(t) = \sup_{1\le j\le d_n}\left|\sum_{r=1}^j \xi(t,r)\right| \quad\text{and}\quad \xi(t,r) = \sqrt{n}\int_0^t\Big(\int_{\mathcal{I}_r}\bar{f}_1\,dx\Big)\frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}.$$

Since $\eta(t)$ is a positive submartingale, Doob's inequality gives $E\sup_t \eta^2(t) \le 4E\eta^2(1)$. Also, since $E\xi(1,r) = 0$, and $E\xi(1,j)\xi(1,k) = 0$ for all $1 \le j \ne k \le d_n$, we can apply Menchoff's inequality (see, e.g., Shorack and Wellner, 1986) here to get

$$\begin{aligned}E\eta^2(1) &\le \Big(\frac{\log 4d_n}{\log 2}\Big)^2\sum_{r=1}^{d_n} E\int_0^1\Big(\int_{\mathcal{I}_r}\bar{f}_1(s,x)\,dx\Big)^2\frac{n\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2}\,ds \\ &\le O(\log d_n)^2\sum_{r=1}^{d_n}\int_0^1 E\Big[\Big(\int_{\mathcal{I}_r}\bar{f}_1(s,x)\,dx\Big)^2\frac{n}{Y_r^{(n)}(s)}\Big]\,ds \\ &\le O(\log d_n)^2\sum_{r=1}^{d_n}\int_0^1 E\Big[\int_{\mathcal{I}_r}\bar{f}_1^2(s,x)\,dx\frac{nw_n}{Y_r^{(n)}(s)}\Big]\,ds \\ &\le O(\log d_n)^2\Big[\sum_{r=1}^{d_n}\int_0^1 E\Big(\int_{\mathcal{I}_r}\bar{f}_1^2(s,x)\,dx\Big)^{\frac{3}{2}}ds\Big]^{\frac{2}{3}}\Big[\sum_{r=1}^{d_n}\int_0^1 E\Big(\frac{nw_n}{Y_r^{(n)}(s)}\Big)^3 ds\Big]^{\frac{1}{3}} \\ &\le O(\log d_n)^2 d_n^{\frac{1}{3}}\Big[\sum_{r=1}^{d_n}\int_0^1 E\Big(w_n^{\frac{1}{2}}\int_{\mathcal{I}_r}|\bar{f}_1(s,x)|^3\,dx\Big)\,ds\Big]^{\frac{2}{3}} \\ &= O(\log d_n)^2\Big[\int_0^1\!\!\int_0^1 E|\bar{f}_1|^3\,ds\,dx\Big]^{\frac{2}{3}}.\end{aligned} \tag{6.14}$$

The second term in (6.13) is bounded by

$$\sqrt{n}\sup_r\int_{\mathcal{I}_r}\sup_t\left|\int_0^t \bar{f}_1\frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}\right|\,dx.$$

which has second moment bounded by

$$
\begin{aligned}
nw_n E \sup_r \int_{\mathcal{I}_r} \sup_t \left| \int_0^t \bar{f}_1 \frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)} \right|^2 dx &\leq nw_n \int_0^1 E \sup_t \left| \int_0^t \bar{f}_1 \frac{M^{(n)}(ds,x)}{Y^{(n)}(s,x)} \right|^2 dx \\
&\leq O(1) \int_0^1\int_0^1 E\left[\bar{f}_1^2 \frac{nw_n}{Y^{(n)}(s,x)}\right] ds\, dx \\
&\leq O(1)\left[\int_0^1\int_0^1 E|\bar{f}_1|^3\, ds\, dx\right]^{\frac{2}{3}}, \qquad (6.15)
\end{aligned}
$$

where Doob's inequality, Hölder's inequality, and (6.8) are used. Therefore

$$
E\|I_1\|^2 \leq O(\log d_n)^2 \left[\int_0^1\int_0^1 E|\bar{f}_1|^3\, ds\, dx\right]^{\frac{2}{3}}. \qquad (6.16)
$$

From (6.16) with $w_n = 1$,

$$
E\|I_2\|^2 = O(1)\left[\int_0^1\int_0^1 E|\bar{f}_1|^2\, ds\, dx\right]^{\frac{2}{3}}. \qquad (6.17)
$$

Next,

$$
\begin{aligned}
E\|I_3\|^2 &\leq \sup_{t,x} E\left[\sqrt{n}\left(\frac{\lambda^{(n)}(s,x)}{Y^{(n)}(s,x)} - \lambda(s,x)\right)\right]^2 \left[\int_0^1\int_0^1 E\bar{f}_1^2\, ds\, dx\right] \\
&= o(1) \int_0^1\int_0^1 E|\bar{f}_1|^2\, ds\, dx, \qquad (6.18)
\end{aligned}
$$

by Lemma 6. Using (6.8) once more,

$$
E\|I_4\|^2 = o(1) \int_0^1\int_0^1 E|\bar{f}_1|^2\, ds\, dx. \qquad (6.19)
$$

It can be checked that

$$
|\bar{f}_1| \leq O(c_n)|\check{h} - h| I(\check{h} \neq 0) + O(1) I(\check{h} = 0),
$$

uniformly in t, z. Thus, since $\check{h} = \hat{h}$ if $\check{h} \neq 0$,

$$
E|\bar{f}_1|^3 \leq O(c_n^3) E|\hat{h} - h|^3 + O(1) P(\check{h} = 0),
$$

so, from Lemmas 4 and 5,

$$
\int_0^1\int_0^1 E|\bar{f}_1|^3\, ds\, dx = O(c_n^3 n^{-\zeta}). \qquad (6.20)
$$

Combining the bounds (6.16–6.20), we find that the second moment of the lhs of (6.10) is of order $O(\log d_n)^2 c_n^2 n^{-\frac{2}{3}\zeta}$, which tends to zero by Condition 3.1. This establishes (6.10).

Step 2 We now prove (6.11). Let

$$
\delta(s,u,x) = \frac{\check{h}^{-1}(s,u)\check{h}^{-\frac{1}{2}}(s,x)}{\int_x^1 \check{h}^{-1}(s,v)\, dv} - \frac{h^{-1}(s,u)h^{-\frac{1}{2}}(s,x)}{\int_x^1 h^{-1}(s,v)\, dv}. \qquad (6.21)
$$

By the arguments of Step 1, the second moment of the lhs of (6.11) is bounded by

$$\begin{aligned}\int_0^\rho E\sup_t\left|\int_0^t\int_x^1 \delta(s,u,x)\,dX(s,u)\right|^2 dx &\leq O(\log d_n)^2\int_0^\rho\left[\int_0^1\int_0^1 E|\delta(s,u,x)|^3\,ds\,du\right]^{\frac{2}{3}} dx\\ &\leq O(\log d_n)^2\left[\int_0^\rho\int_0^1\int_0^1 E|\delta(s,u,x)|^3\,ds\,du\,dx\right]^{\frac{2}{3}}\\ &= O(\log d_n)^2 c_n^5 n^{-\frac{2}{3}\zeta}\to 0,\end{aligned}$$

where the bound on the triple integral is from Lemma 7. □

Proof of Theorem 3.3 Define

$$\mathcal{A}^*(t,z) = z\int_0^t\int_0^1 \lambda(s,x)f(s,x)\,dx\,ds,$$

to which $\bar{\mathcal{A}}(t,z)$ converges in probability under the general alternative. From the definition of $\bar{J}$ in (3.7), it is easily checked that

$$\frac{\partial^2\bar{J}(\mathcal{A}-\mathcal{A}^*)}{\partial t\partial z} = h^{-\frac{1}{2}}(t,z)\left[\lambda(t,z)-\frac{\int_z^1 h^{-1}(t,u)\lambda(t,u)\,du}{\int_z^1 h^{-1}(t,v)\,dv}\right].$$

Suppose that $\bar{J}(\mathcal{A}-\mathcal{A}^*)=0$. Then the expression inside the square brackets above vanishes, so that

$$\lambda(t,z)\int_z^1 h^{-1}(t,v)\,dv = \int_z^1 h^{-1}(t,u)\lambda(t,u)\,du.$$

Taking partial derivatives wrt z both sides gives that $\partial\lambda(t,z)/\partial z = 0$ for $(t,z)\in[0,1]\times[0,\rho]$, so that H_0 holds, contrary to the premise of the theorem. Thus, $\bar{J}(\mathcal{A}-\mathcal{A}^*))\neq 0$. From arguments in the proof of Theorem 3.2, it can be seen that $\|(\hat{J}-\bar{J})(\mathcal{A}-\mathcal{A}^*)\|\xrightarrow{P}0$. Hence,

$$\|\hat{J}(\mathcal{A}-\mathcal{A}^*)\|\xrightarrow{P}\|\bar{J}(\mathcal{A}-\mathcal{A}^*)\|>0. \tag{6.22}$$

Along the lines of the proof of Theorem 3.2, it can be shown that $\hat{J}(\sqrt{n}(\bar{\mathcal{A}}-\mathcal{A}^*))$ converges weakly in $D_2([0,1]\times[0,\rho])$, although not necessarily to Brownian sheet, cf. the proof of Proposition 4.3 of MU. Similarly, using MU (Proposition 3.2), it can be shown that $\hat{J}(\sqrt{n}(\hat{\mathcal{A}}-\mathcal{A}))$ converges in the same sense. The triangle inequality gives

$$\sqrt{n}\|\hat{J}(\mathcal{A}-\mathcal{A}^*)\|\leq\|\hat{J}(\sqrt{n}(\hat{\mathcal{A}}-\mathcal{A}))\|+S+\|\hat{J}(\sqrt{n}(\bar{\mathcal{A}}-\mathcal{A}^*))\| = S+O_P(1),$$

the equality holding by the continuous mapping theorem, so that $S\xrightarrow{P}\infty$ by (6.22). Thus the test is consistent. □

Appendix

The following lemma is routine.

Lemma 1 *Let h, h_1, h_2 be functions on $[0,1]^2$ that have bounded variation and are Lipschitz, with h nonnegative and bounded away from zero. Then $1/h$, $\sqrt{h_1}$ and h_1h_2 have bounded variation and are Lipschitz.*

The next two lemmas collect some properties of weak integrals in the plane. The first is a version of the integration by parts formula. Let $\phi : \mathcal{J} \to I\!R$, where $\mathcal{J} = [a, a'] \times [b, b']$.

Lemma 2 *If $\xi : \mathcal{J} \to \boldsymbol{R}$ is continuous and $\phi, \phi(a,\cdot)$, and $\phi(\cdot, b)$ have bounded variation, then the weak integral $\iint_{\mathcal{J}} \phi \, d\xi$ exists and is equal to:*

$$\iint_{\mathcal{J}} d(\phi(s,x)\xi(s,x)) + \iint_{\mathcal{J}} \xi(s,x)\, d\phi(s,x) - \int_a^{a'} \xi(s,b')\, d\phi(s,b')$$
$$+ \int_a^{a'} \xi(s,b)\, d\phi(s,b) - \int_b^{b'} \xi(a',x)\, d\phi(a',x) + \int_b^{b'} \xi(a,x)\, d\phi(a,x). \tag{A.1}$$

Proof Theorem 9.3 of Hildebrandt (1963) gives that the weak integral

$$\iint_{\mathcal{J}} (\xi(s,x) - \xi(s,b) - \xi(a,x) + \xi(a,b))\, d\phi(s,x) \tag{A.2}$$

exists, and coincides with the weak integral

$$\iint_{\mathcal{J}} (\phi(s,x) - \phi(s,b') - \phi(a',x) + \phi(a',b'))\, d\xi(s,x), \tag{A.3}$$

which exists by Theorem 8.8 of Hildebrandt. Theorem 5.8 of Hildebrandt shows that $\phi(\cdot, x)$ and $\phi(s,\cdot)$ are of bounded variation for fixed s, x. (A.1) can then be obtained by rearranging the terms in (A.2) and (A.3). □

Lemma 3 *Let ξ be a stochastic process on $\mathcal{J}$. If the weak integral $\iint_{\mathcal{J}} \phi \, d\xi$ exists a.s., and the stochastic integral $\iint_{\mathcal{J}} \phi \, d\xi$ exists in the L^2-sense, then they coincide a.s.*

Proof The result follows immediately from the definitions of the stochastic integral and the weak integral, and the fact that an L^2-limit agrees almost surely with an a.s.-limit. □

The next lemma is a refined version of Proposition 3.3 of MU, giving a *rate* of convergence of $\hat{h}$ to h.

Lemma 4 *There exists $\zeta > 0$ such that*

$$\int_0^1\int_0^1 E|\hat{h}(t,z) - h(t,z)|^3\, dt\, dz = O(n^{-\zeta}).$$

Proof We shall use much of the notation of MU (proof of Proposition 3.3), without re-defining it here. As in MU,

$$|\hat{h} - h|^3 \le O(1)[|\hat{h} - \tilde{h}|^3 + |h - h^0|^3 + |h^0 - h^\dagger|^3 + |h^\dagger - h^*|^3 + |R|^3]. \tag{A.4}$$

For the first term,

$$\begin{aligned}\sup_{t,z} |\hat{h}(t,z) - \tilde{h}(t,z)|^3 &\le \sup_{t,z} \left[\frac{1}{b_n^2} \sum_{j,r} K\left(\frac{t-\tau_j}{b_n}\right) \left| K\left(\frac{z-x_r}{b_n}\right) - \frac{1}{w_n}\int_{I_r} K\left(\frac{z-x}{b_n}\right) dx \right| \Delta_{jr}\right]^3 \\ &= O(w_n^3 b_n^{-9}) \hat{H}^3(1,1). \end{aligned} \tag{A.5}$$

Also,

$$E\tilde{H}^3(1,1) = E\Big[nw_n \int_0^1\!\!\int_0^1 \frac{N^{(n)}(ds,x)}{(Y^{(n)}(s,x))^2}\,dx\Big]^3$$

$$\le 8E\Big[nw_n \int_0^1\!\!\int_0^1 \frac{\lambda^{(n)}(s,x)}{(Y^{(n)}(s,x))^2}\,ds\,dx\Big]^3 + 8(nw_n)^3 E\Big|\int_0^1\!\!\int_0^1 \frac{M^{(n)}(ds,x)}{(Y^{(n)}(s,x))^2}\,dx\Big|^3. \tag{A.6}$$

The first term in (A.6) can be shown to be bounded using (6.8). The expectation in the second term in (A.6) is bounded by

$$\int_0^1 E\Big|\int_0^1 \frac{M^{(n)}(ds,x)}{(Y^{(n)}(s,x))^2}\Big|^3 dx \le \int_0^1 \left(EM_1^4(1)\right)^{\frac{3}{4}} dx. \tag{A.7}$$

where

$$M_1(t) = \int_0^t \frac{M^{(n)}(ds,x)}{(Y^{(n)}(s,x))^2}$$

and we have suppressed the dependence of $M_1(t)$ on x. Let $[M_1]$ and $\langle M_1\rangle$ be the quadratic variation and the predictable quadratic variation of martingale M_1, respectively. We shall use the Burkholder-Davis-Gundy inequality (Dellacherie and Meyer, 1982, p.287)

$$E\sup_{v\in[0,t]} M_1^4(v) \le CE[M_1]_t^2. \tag{A.8}$$

Since the square integrable martingale M_1 is of integrable variation, it has no continuous part. Hence $[M_1]_t = \sum_{v\le t}(\Delta M_1(v))^2$. The process

$$\xi(t) = [M_1]_t - \langle M_1\rangle_t = \int_0^t \frac{M^{(n)}(ds,x)}{(Y^{(n)}(s,x))^4}$$

is a martingale, so $E[M_1]_1^2 \le 2E\langle M_1\rangle_1^2 + 2E\xi^2(1) = 2E\langle M_1\rangle_1^2 + 2E\langle\xi\rangle_1$. But, by (6.8),

$$\begin{aligned} E\langle M_1\rangle_1^2 &= E\Big[\int_0^1 \frac{\lambda^{(n)}(s,x)}{(Y^{(n)}(s,x))^4}\,ds\Big]^2 = O\left(\frac{1}{nw_n}\right)^6, \\ E\langle\xi\rangle_1 &= E\int_0^1 \frac{\lambda^{(n)}(s,x)}{(Y^{(n)}(s,x))^8}\,ds = O\left(\frac{1}{nw_n}\right)^7. \end{aligned}$$

It then follows from (A.7) and (A.8) that the second term in (A.6) is of order

$$O(nw_n)^3 O\left(\frac{1}{nw_n}\right)^{6\cdot\frac{3}{4}} \to 0,$$

so $E\tilde{H}^3(1,1) < \infty$, and from (A.5),

$$E\sup_{t,z} |\hat{h}(t,z) - \tilde{h}(t,z)|^3 = O(w_n^3 b_n^{-9}). \tag{A.9}$$

Since $\sup_{t,z \geq b_n} |h(t,z) - h^0(t,z)| = O(b_n)$, by the Lipschitz condition on h,

$$\int_0^1 \int_0^1 |h(t,z) - h^0(t,z)|^3 \, dt \, dz = O(b_n). \tag{A.10}$$

For the third term in (A.4),

$$\begin{aligned}
&\sup_{t,z} |h^0(t,z) - h^\dagger(t,z)| \\
&= \sup_{t,z} \left| \frac{1}{b_n^2} \int_0^1 K\left(\frac{t-s}{b_n}\right) \left[\int_0^1 K\left(\frac{z-x}{b_n}\right) h(s,x) \, dx - \frac{1}{d_n} \sum_{r=1}^{d_n} K\left(\frac{z-x_r}{b_n}\right) h(s,x_r) \right] ds \right| \\
&\leq \frac{1}{b_n} \sup_{s,z} \left[\sum_{r=1}^{d_n} \int_{I_r} \left| K\left(\frac{z-x}{b_n}\right) h(s,x) - K\left(\frac{z-x_r}{b_n}\right) h(s,x_r) \right| dx \right. \\
&\leq \frac{1}{b_n} \sup_{s,z} \left[\sum_{r=1}^{d_n} \int_{I_r} \left| K\left(\frac{z-x}{b_n}\right) - K\left(\frac{z-x_r}{b_n}\right) \right| h(s,x_r) \, dx \right] + O(w_n) \\
&\leq \frac{1}{b_n} O(w_n b_n^{-1}) + O(w_n) = O(w_n b_n^{-2}).
\end{aligned} \tag{A.11}$$

For the fourth term in (A.4),

$$\begin{aligned}
&\sup_{t,z} E|h^\dagger(t,z) - h^*(t,z)|^3 \\
&= \sup_{t,z} E\left| \frac{1}{b_n^2 d_n} \sum_{r=1}^{d_n} K\left(\frac{z-x_r}{b_n}\right) \int_0^1 K\left(\frac{t-s}{b_n}\right) \left(h(s,x_r) - n w_n \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} \right) ds \right|^3 \\
&\leq \left(\frac{1}{b_n^2 d_n}\right)^3 \sup_{t,z} \left[\sum_{r=1}^{d_n} K^{\frac{3}{2}}\left(\frac{z-x_r}{b_n}\right) \right]^{\frac{2}{3} \cdot 3} \\
&\qquad \times E\left[\sum_{r=1}^{d_n} \left| \int_0^1 K\left(\frac{t-s}{b_n}\right) \left(h(s,x_r) - n w_n \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} \right) ds \right|^3 \right]^{\frac{1}{3} \cdot 3} \\
&\leq O\left(\frac{1}{b_n^2 d_n}\right)^3 (b_n d_n)^2 \sup_{t,z} E\left[\sum_{r=1}^{d_n} \left| \int_0^1 K\left(\frac{t-s}{b_n}\right) \left(h(s,x_r) - n w_n \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} \right) ds \right|^3 \right] \\
&\leq O\left(\frac{1}{b_n^2 d_n}\right)^3 (b_n d_n)^2 \sup_{t,z} \sum_{r=1}^{d_n} \int_0^1 K^3\left(\frac{t-s}{b_n}\right) E\left| h(s,x_r) - n w_n \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} \right|^3 ds \\
&\leq O\left(\frac{1}{b_n^2 d_n}\right)^3 (b_n d_n)^3 \sup_{s,r} E\left| h(s,x_r) - n w_n \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} \right|^3.
\end{aligned}$$

Using the Lipschitz property of λ and (6.8), the supremum term above is bounded by

$$\begin{aligned}
&O(w_n^3) + O(1) \sup_{s,r} E\left| \frac{Y_r^{(n)}(s)}{n w_n} - f(s,x_r) \right|^3 + O(1) \sup_{s,r} P(Y_r^{(n)}(s) = 0) \\
&\leq O(w_n^3) + O(1) \left[O\left(\frac{1}{n w_n}\right)^2 \right]^{\frac{3}{4}} + O(e^{-C n w_n}) = O\left(\frac{1}{n w_n}\right)^{\frac{3}{2}}.
\end{aligned}$$

where we have used the Lipschitz property of f and the fact that $Y_r^{(n)}(s)$ is a binomial r.v. with mean of order $O(nw_n)$ to bound its fourth central moment. Therefore

$$\sup_{t,z} E|h^\dagger(t,z) - h^*(t,z)|^3 = O\left(\frac{1}{b_n^3}\right) O\left(\frac{1}{nw_n}\right)^{\frac{3}{2}} = O\left(\frac{1}{nw_n b_n^2}\right)^{\frac{3}{2}}. \tag{A.12}$$

Finally, for the fifth term in (A.4),

$$\sup_{t,z} E|R(t,z)|^3 \le \sup_{t,z}[ER^4(t,z)]^{\frac{3}{4}}. \tag{A.13}$$

Let

$$R(t,z = \frac{n}{b_n^2 d_n^2} \sum_{r=1}^{d_n} K\left(\frac{z-x_r}{b_n}\right) \int_0^u K\left(\frac{t-s}{b_n}\right) \frac{dM_r^{(n)}(s)}{(Y_r^{(n)}(s))^2}.$$

Then $R(t,z,\cdot)$ is a m rtingale. Using the same arguments that were applied to M_1,

$$ER^4(t,z) \le E \sup_{0\le u\le 1} R^4(t,z,u) \le CE[R(t,z,\cdot)]_1^2 \le O(1)(E\langle R(t,z,\cdot)\rangle_1^2 + E\xi^2(1)). \tag{A.14}$$

where $\xi(u) = [R(t,z,\cdot)]_u - \langle R(t,z,\cdot)\rangle_u$ is a martingale. Since no two of the counting processes $N_r^{(n)}$, $r = 1, 2, \cdots, d_n$ jump simultaneously,

$$\begin{aligned}
[R(t,z,\cdot)]_u &= \sum_{v\le u} (\Delta R(t,z,v))^2 \\
&= \sum_{v\le u} \sum_{r=1}^{d_n} \left(\frac{n}{b_n^2 d_n^2}\right)^2 K^2\left(\frac{z-x_r}{b_n}\right) K^2\left(\frac{t-v}{b_n}\right) \frac{\Delta N_r^{(n)}(v)}{(Y_r^{(n)}(v))^4} \\
&= \left(\frac{n}{b_n^2 d_n^2}\right)^2 \sum_{r=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) \int_0^u K^2\left(\frac{t-s}{b_n}\right) \frac{dN_r^{(n)}(s)}{(Y_r^{(n)}(s))^4}.
\end{aligned}$$

It follows that

$$\xi(u) = \left(\frac{n}{b_n^2 d_n^2}\right)^2 \sum_{r=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) \int_0^u K^2\left(\frac{t-s}{b_n}\right) \frac{dM_r^{(n)}(s)}{(Y_r^{(n)}(s))^4}. \tag{A.15}$$

The last two terms in (A.14) can be bounded above as follows:

$$\begin{aligned}
\sup_{t,z} E\langle R(t,z,\cdot)\rangle_1^2 &= \left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} E\left[\sum_{r=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) \int_0^1 K^2\left(\frac{t-s}{b_n}\right) \frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^4}\, ds\right]^2 \\
&\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-6} \sup_{t,z} \sum_{r,l=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) K^2\left(\frac{z-x_l}{b_n}\right) \\
&\quad \times \int_0^1\int_0^1 K^2\left(\frac{t-s}{b_n}\right) K^2\left(\frac{t-v}{b_n}\right) ds\, dv \\
&\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-6} b_n^2 (b_n d_n)^2 \\
&= O(1)(n^2 b_n^4)^{-1}.
\end{aligned} \tag{A.16}$$

$$
\begin{aligned}
\sup_{t,z} E\xi^2(1) &= \sup_{t,z} E\langle\xi\rangle_1 \\
&= \left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} \sum_{r=1}^{d_n} K^4\left(\frac{z-x_r}{b_n}\right) \int_0^1 K^4\left(\frac{t-s}{b_n}\right) E\frac{\lambda_r^{(n)}(s)}{(Y_r^{(n)}(s))^8}\, ds \\
&\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} \sum_{r=1}^{d_n} K^4\left(\frac{z-x_r}{b_n}\right) \int_0^1 K^4\left(\frac{t-s}{b_n}\right) E(Y_r^{(n)}(s))^{-7}\, ds \\
&\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-7}(b_n d_n) b_n \\
&= O(1)(n^3 b_n^6)^{-1}. \qquad (A.17)
\end{aligned}
$$

From (A.13–A.17), we get

$$
\sup_{t,z} E|R(t,z)|^3 \le O(1)[n^2 b_n^4]^{-\frac{3}{4}}. \qquad (A.18)
$$

Finally combining the bounds for the five terms in (A.4), we have

$$
\begin{aligned}
\int_0^1\int_0^1 E\left|\hat{h}(t,z) - h(t,z)\right|^3 dt\, dz &= O[w_n^3 b_n^{-9} + b_n + w_n^3 b_n^{-6} + (nw_n b_n^2)^{-\frac{3}{2}} + (nb_n^2)^{-\frac{3}{2}}] \\
&= O(n^{-\zeta})
\end{aligned}
$$

by Condition 3.1, where $\zeta = \min(\beta, \frac{51}{8}(\alpha - \frac{9}{17})) > 0$. □

Lemma 5

$$
\int_0^1\int_0^1 P(\check{h}(t,z) = 0)\, dt\, dz = O(n^{-\zeta}).
$$

Proof Let $0 < c \le C < \infty$ be lower and upper bounds for h.

$$
\begin{aligned}
P(\check{h}(t,z) = 0) &= P(\hat{h}(t,z) \le c_n^{-1}) + P(\hat{h}(t,z) \ge c_n) \\
&\le P(\hat{h}(t,z) - h(t,z) \le c_n^{-1} - c) + P(\hat{h}(t,z) - h(t,z) \ge c_n - C) \\
&= P(h(t,z) - \hat{h}(t,z) \ge c - c_n^{-1}) + P(\hat{h}(t,z) - h(t,z) \ge c_n - C) \\
&\le \frac{E\left|\hat{h}(t,z) - h(t,z)\right|^3}{(c - c_n^{-1})^3} + \frac{E\left|\hat{h}(t,z) - h(t,z)\right|^3}{(c_n - C)^3}
\end{aligned}
$$

so the result follows by Lemma 4. □

Lemma 6

$$
\sup_{t,z} E\left[\sqrt{n}\left(\frac{\lambda^{(n)}(t,z)}{Y^{(n)}(t,z)} - \lambda(t,z)\right)\right]^2 \to 0.
$$

Proof The expression on the lhs above is bounded by

$$
\begin{aligned}
&\sup_{t,z} E\left[\sqrt{n}O(w_n)I(Y^{(n)}(t,z) \neq 0) + O(\sqrt{n})I(Y^{(n)}(t,z) = 0)\right]^2 \\
&\le O(\sqrt{nw_n^2}) + O(n)\sup_{t,z} P(Y^{(n)}(t,z) = 0) \\
&\le O(n^{\frac{1}{2}(1-2\alpha)}) + O(ne^{-Cn^{1-\alpha}}) \to 0,
\end{aligned}
$$

where the last inequality comes from (6.9). □

Lemma 7 *For $\delta(s,u,x)$ defined by (6.21),*

$$\int_0^\rho\int_0^1\int_0^1 E|\delta(s,u,x)|^3\,ds\,du\,dx = O(c_n^{7.5}n^{-\zeta}).$$

Proof Let

$$I_1 = I(\check{h}(s,u)\neq 0,\ \check{h}(s,x)\neq 0,\ \int_x^1 \check{h}^{-1}(s,v)\,dv\neq 0),$$
$$I_2 = I(\check{h}(s,u)=0) + I(\check{h}(s,x)=0) + I(\int_x^1 \check{h}^{-1}(s,v)\,dv = 0).$$

Then

$$\begin{aligned}
|\delta(s,u,x)| &= O(c_n)\Big|(\check{h}^{-1}(s,u)-h^{-1}(s,u))\check{h}^{-\frac{1}{2}}(s,x)\int_x^1 h^{-1}(s,v)\,dv \\
&\quad +(\check{h}^{-\frac{1}{2}}(s,x)-h^{-\frac{1}{2}}(s,x))h^{-1}(s,u)\int_x^1 h^{-1}(s,v)\,dv \\
&\quad +h^{-\frac{1}{2}}(s,x)h^{-1}(s,u)\int_x^1(h^{-1}(s,v)-\check{h}^{-1}(s,v))\,dv\Big|I_1 + O(I_2) \\
&= O(c_n)\Big[c_n^{\frac{1}{2}}|\check{h}^{-1}(s,u)-h^{-1}(s,u)| + |\check{h}^{-\frac{1}{2}}(s,x)-h^{-\frac{1}{2}}(s,x)| \\
&\quad +\int_x^1|h^{-1}(s,v)-\check{h}^{-1}(s,v)|\,dv\Big]I_1 + O(I_2) \\
&= O(c_n)\Big[c_n^{\frac{3}{2}}|\hat{h}(s,u)-h(s,u)| + c_n|\hat{h}(s,x)-h(s,x)| \\
&\quad +c_n\int_x^1|\hat{h}(s,v)-h(s,v)|I(\check{h}(s,v)\neq 0)\,dv \\
&\quad +\int_x^1 I(\check{h}(s,v)=0)\,dv\Big] + O(I_2).
\end{aligned}$$

where we have used the fact that $\check{h}=\hat{h}$ when $\check{h}\neq 0$. Thus,

$$\begin{aligned}
E|\delta(s,u,x)|^3 \;&\leq\; O(c_n)^3\Big[c_n^{\frac{9}{2}}E|\hat{h}(s,u)-h(s,u)|^3 + c_n^3 E|\hat{h}(s,x)-h(s,x)|^3 \\
&\quad +c_n^3\int_0^1 E|\hat{h}(s,v)-h(s,v)|^3\,dv + \int_0^1 P(\check{h}(s,v)=0)\,dv\Big] \\
&\quad +O(1)\Big[P(\check{h}(s,u)=0)+P(\check{h}(s,x)=0)+P\Big(\int_x^1\check{h}^{-1}(s,v)\,dv=0\Big)\Big].
\end{aligned}$$

Notice that if $\check{h}(s,1)\neq 0$, then $c_n^{-1}<\hat{h}(s,1)<c_n$. Since $\hat{h}(s,\cdot)$ is continuous, $c_n^{-1}<\hat{h}(s,v)<c_n$ in a small interval $v\in[1-\varepsilon,1]$, $\varepsilon>0$. Hence $c_n^{-1}<\check{h}^{-1}(s,v)<c_n$, for $v\in[1-\varepsilon,1]$. Hence $\int_x^1\check{h}^{-1}(s,v)\,dv\neq 0$, for $x<1$. Therefore,

$$P\Big(\int_x^1\check{h}^{-1}(s,v)\,dv=0\Big)\leq P(\check{h}(s,1)=0),\qquad \text{for } x<1.$$

Now apply Lemmas 4, 5 and the fact that $\int_0^1 P(\check{h}^{-1}(t,1)=0)\,dt = O(n^{-\zeta})$, which is easy to see from proofs of Lemmas 4 and 5, to complete the proof.

Acknowledgments

The authors thank Estate Khmaladze for his encouragement, Jayaram Sethuraman for help in refining the proof of Theorem 2.1, and Peter Sasieni for furnishing them with data from the British Medical Research Council's 4th myelomatosis trial.

References

Aalen, O. O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 701–726.

Adler, R. J. (1991). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes.* Inst. Math. Statist.

Andersen, P. K. Borgan, O. Gill, R. D. and Keiding, N. (1982). Linear non-parametric tests for comparison of counting processes, with applications to censored survival data (with discussion). *Int. Statist. Rev.* **50** 219–258. Correction **52** 225.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B.* **34** 187–220.

Cuzick, J. Cooper, E. H. and MacLennan, I. C. M. (1985). The prognostic value of serum β_2 microglobulin compared with other presentation features in myelomatosis. *British J. Cancer* **52** 1–6.

Dellacherie, C. and Meyer, P. A. (1982). *Probabilities and Potential B.* North Holland.

Doob, J. (1949). Heurstic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **2** 279–290.

Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of *P*. *J. Roy. Statist. Soc.* **85** 87–94.

Fisher, R. A. (1924). The conditions under which χ^2 measures the discrepancy between observation and hypothesis. *J. Roy. Statist. Soc.* **87** 442–450.

Hildebrandt, T. H. (1963). *Introduction to the Theory of Integration.* Academic Press.

Hjort, N. L. (1990). Goodness of fit tests in models for life history data based on cumulative hazard rates. *Ann. Statist.* **18** 1221–1258.

Keiding, N. and Gill D. (1990). Random truncation models and Markov processes. *Ann. Statist.* **18** 582–602

Khmaladze, E. V. (1981). Martingale approach to the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26** 246–265.

Liptser, R. S. and Shiryayev, A. N. (1977). *Statistics of Random Processes I*. Springer-Verlag.

McKeague, I. W. and Utikal, K. J. (1990). Identifying nonlinear covariate effects in semimartingale regression models. *Probab. Th. Rel. Fields* **87** 1–25.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

Nikabadze, A. M. (1987). On a method for constructing goodness-of-fit tests for parametric hypotheses in $\mathbb{R}^m$. *Theory Probab. Appl.* **32** 539–544.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil. Mag. Ser. 5* **50** 157–175.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley.

Wong, E. and Zakai, M. (1974). Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrsch. verw. Gebiete* **29** 109–122.